

J. Perinat. Med.  
2 (1974) 110

## Computer prediction of neonatal conditions by sampled data analyses of the fetal heart rate and the scalp blood during labor

H. Takemura, H. Fujimoto

Department of Obstetrics and Gynecology, Osaka University Medical School, Osaka, Japan

Shionogi Pharmaceutical Company, Computer Center, Osaka, Japan

Received January 12, 1973. Accepted December 18, 1973.

One of the most important pieces of information that an obstetrician can get, as feedback in the attendance of childbirths, is the expected neonatal condition as represented by the APGAR score [1]. To evaluate the quality of labor monitoring, some guidance is needed during the course of labor. This cannot be achieved without such type of information as that which is usually called a controlled variable in control engineering. Although its essential significance as to the later development of the baby's body and mind is somewhat controversial, a collaborative study reported by DRÄGE, KENNEDY, BERENDES, SCHWARZ and WEISS [3] has clearly demonstrated that the lower the points of the APGAR score, the higher the incidence of neurological abnormalities at one year of age. It is also commonly known that the high mortality and morbidity rates are in the low APGAR score group. Hence, in order to establish a mathematical formula for predicting the APGAR scores at both one and five minutes after birth from several fetal parameters minimizing any possibility of personal, subjectively biased judgments, two different types of computer prediction systems have been developed. One is **discriminant analysis** of fetal variables to differentiate between the normal infants with an APGAR score of 8 or more and the depressed infants with a score of 7 or less. The other is **partial regression analysis** of fetal variables to construct a linear

### Curriculum vitae

HARUMI FUJIMOTO, B. S.  
*Born in Saitama, Japan on April 24, 1924, he graduated from Tokyo University of Education in 1965 with the degree of B. S., majoring Mathematics in the School of Science. Presently he is working as a system analyst in Shionogi Computing Center, Shionogi Research Laboratory, Osaka, Japan.*



regression function of APGAR scores using a set of data which were individually given at each stage of labor.

### 1. Materials and methods

More than one hundred cases who were admitted at Yale New Haven Medical Center, were continuously monitored by electronic means as well as by fetal scalp blood analysis during the year 1966—67. Only the data of the cases with complete sets of information such as fetal blood pH, PO<sub>2</sub>, PCO<sub>2</sub>, actual base excess (BE), the level of basal fetal heart rate (BFHR) and the incidences of late and variable fetal heart rate decelerations (%UPI & % CORD) during twenty minutes at each stage of labor, were processed by an electronic digital computer FACOM 270—20\*.

\* Fuji Electric Co., Japan.

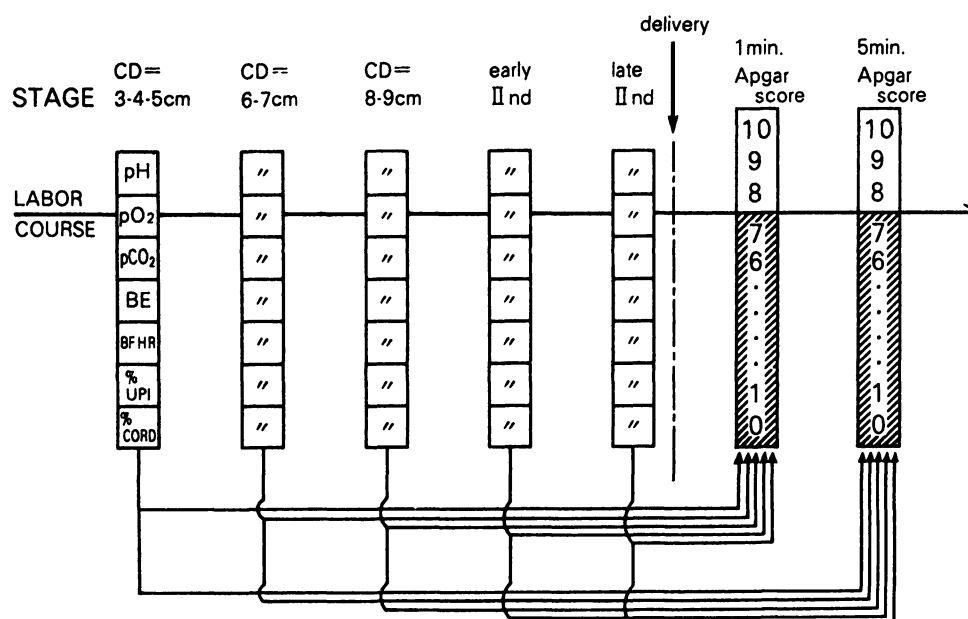


Fig. 1. Design of a sampled data prediction system for neonatal conditions based upon a set of seven fetal parameters at a given stage of childbirth.

Tab. I. The number of subjects studied in this report.

stage	First Stage			Second Stage	
	3—4—5	6—7	8—9	early	late
No.	47	22	12	14	58

The number of cases computed is shown in Tab. I.

As shown in Fig. 1, a course of labor was divided into five stages; the early first stage where cervical dilatation was 3—4—5 cm, the middle first stage of 6—7 cm, the late first of 8—9 cm, the early second stage, and the late second stage which was the last twenty minutes just before delivery. At each stage the quantitative data for processing were obtained during a period of twenty minutes. Several measurements of fetal blood, analysed for pH with an ASTRUP pH-meter and for PO<sub>2</sub> and PCO<sub>2</sub> with a micro IL meter during the twenty minutes were all averaged to get a set of sampled data at a given stage for processing. Actual base excess (BE) was calculated with the SIGGAARD-ANDERSEN alignment nomogram [8] from the averaged values of pH, PCO<sub>2</sub> and hemoglobin concentration. The basal fetal heart rate (BFHR as an initial value of fetal heart rate just before each contraction) was

also averaged as all uterine contractions during the twenty minutes to get the level of BFHR at that stage. The parameters in the pattern of late FHR deceleration (UPI) and that of variable one (CORD) were both represented by the ratio of their incidence or appearance over the number of contractions during the twenty minutes. The identification of the FHR patterns was only possible visually [4]. APGAR scores at one and five minutes after birth were correlated with a set of values of seven fetal parameters mentioned above, at each stage of labor, in order to get a predictive solution by both discriminant analysis and partial regression analysis.

## 2. Results

### 2.1 Discriminant analysis

A computing flow chart for predictive discrimination between the vigorous infants who are assumed to have 8 or more APGAR score and the depressed one of 7 or less at one and five minutes after birth, in any stage of labor, is shown in Fig. 2. The sets of data that were provided to the computer (from forty-seven patients at the early first stage) are given in Tab. II. The discriminant functions at all the five stages for both one and five minutes APGAR scores, which thus obtained, are summarized in Fig. 3.

Tab. II. Examples of complete set data of forty-seven fetuses at the early first stage of labor with 3—4—5 cm cervical dilatation.

No.	Case No.	Actual APGAR s.	pH	PO <sub>2</sub>	PCO <sub>2</sub>	BE	Mean BFHR	%UPI per cent.	%CORD per cent.
1	21	9.00	7.31	15.00	41.00	-4.79	155.00	1.00	1.00
2	24	9.00	7.31	21.99	44.50	-3.69	169.99	0.71	0.28
3	25	9.00	7.28	24.00	43.29	-6.00	140.00	0.18	0.45
4	26	8.00	7.21	24.39	50.00	-4.79	140.00	0.14	0.42
5	28	1.00	7.26	16.69	56.00	-3.19	180.00	0.60	0.00
6	29	9.00	7.35	20.00	42.50	-1.89	140.00	0.00	1.00
7	30	9.00	7.28	23.00	51.79	-2.69	140.00	0.00	0.00
8	31	8.00	7.23	13.99	56.50	-4.69	150.00	0.83	0.50
9	32	9.00	7.32	25.00	51.29	0.00	144.99	0.00	0.33
10	33	7.00	7.26	20.50	56.50	-3.09	135.00	0.28	0.42
11	36	10.00	7.31	18.50	50.00	-1.29	150.00	0.00	0.00
12	37	9.00	7.27	20.50	49.69	-3.79	150.00	0.00	0.00
13	38	8.00	7.29	21.19	44.79	-4.39	130.00	0.00	0.29
14	39	8.00	7.31	20.50	45.29	-3.19	130.00	0.27	0.45
15	40	7.00	7.26	12.79	53.69	-3.09	144.99	0.62	0.50
16	41	9.00	7.29	20.29	46.89	-4.19	144.99	0.00	0.09
17	44	8.00	7.01	12.79	59.29	-17.09	105.00	0.66	0.33
18	46	1.00	7.26	16.79	49.69	-5.00	144.99	0.12	0.62
19	48	9.00	7.28	15.50	52.29	-3.19	155.00	0.25	0.25
20	49	8.00	7.29	21.99	42.00	-5.29	130.00	0.00	0.00
21	50	9.00	7.29	15.00	47.50	-3.39	160.00	0.00	0.37
22	53	9.00	7.26	17.50	49.00	-5.50	160.00	0.60	0.20
23	56	2.00	7.32	11.49	54.00	0.50	144.99	0.25	0.00
24	57	6.00	7.31	18.99	49.00	-1.59	140.00	0.00	0.66
25	60	5.00	7.21	10.39	54.59	-6.59	160.00	1.00	0.66
26	61	7.00	7.27	25.00	42.50	-6.79	140.00	0.00	0.30
27	62	9.00	7.37	16.00	42.29	-0.09	150.00	0.12	0.62
28	64	5.00	7.31	19.69	41.50	-5.79	135.00	0.62	0.00
29	71	9.00	7.26	14.69	44.69	-6.79	150.00	0.18	0.45
30	72	8.00	7.25	17.19	55.50	-4.39	150.00	0.20	0.46
31	73	8.00	7.28	23.69	42.69	-6.00	130.00	0.00	0.37
32	74	9.00	7.28	18.29	50.59	-3.29	150.00	0.00	0.00
33	75	6.00	7.33	25.00	42.00	-3.24	173.00	0.00	0.14
34	78	8.00	7.30	16.49	43.00	-4.79	140.00	0.00	0.53
35	80	3.00	7.30	18.00	47.69	-3.00	135.00	0.00	0.90
36	81	2.00	7.26	10.79	46.69	-6.29	110.00	0.42	0.00
37	82	9.00	7.31	20.29	52.79	0.29	150.00	0.12	0.75
38	83	9.00	7.28	22.29	42.29	-6.09	150.00	0.00	0.50
39	84	8.00	7.29	14.69	46.29	-3.39	130.00	0.00	0.00
40	88	3.00	7.31	21.29	47.00	-2.79	130.00	0.00	0.53
41	92	9.00	7.25	20.50	46.19	-6.69	150.00	0.07	0.35
42	93	8.00	7.31	19.49	41.00	-4.89	169.99	0.00	0.33
43	96	9.00	7.33	20.29	49.50	0.00	164.99	0.00	0.29
44	97	9.00	7.29	21.99	45.09	-4.39	130.00	0.00	0.42
45	98	9.00	7.29	20.29	46.00	-4.00	140.00	0.14	0.71
46	99	7.00	7.33	24.49	42.59	-2.89	155.00	0.00	0.00
47	100	9.00	7.21	20.00	35.50	-6.19	164.99	0.50	0.60

K: number of variable  
 N(1): sample size from population 1  
 N(2): sample size from population 2  
 AP: probability that an observation comes from population 1  
 AQ: probability that an observation comes from population 2

V(i,j): covariance matrix  
 A(i,j): mean of X(j) of samples from population i  
 C(i,j): the cost of misclassification  
 X(i): observation, i=1,...,K

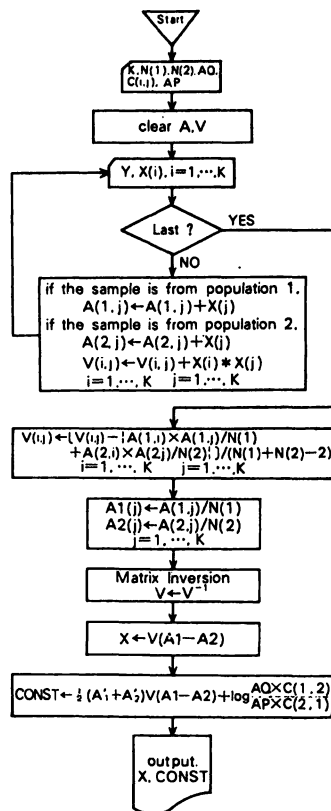
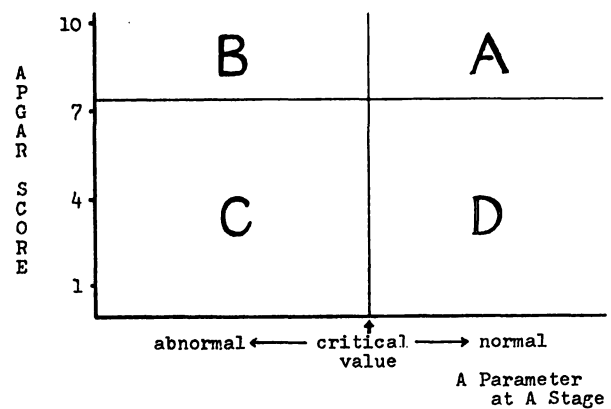


Fig. 2. Computing flow chart for discriminant analysis of the neonatal conditions based upon several fetal parameters.

The functions are constructed so as to have a positive value if the infant is to be vigorous and to have a negative one if it is to be depressed. The correct and false discriminations of neonatal



#### DIAGNOSTIC PROBABILITY OF LABOR MONITORING

O. Relative Differential Probability for Correct Diagnosis

$$= \frac{(A+C)/(A+B+C+D)}{(A+B)/(A+B+C+D)} - 1 = \frac{C-B}{A+B}$$

I. Probability for Correct Diagnosis =  $\frac{A+C}{A+B+C+D}$

II. Detectability for The Depressed =  $\frac{C}{C+D}$

III. Detectability for The Vigorous =  $\frac{A}{A+B}$

IV. Accuracy as for The Depressed =  $\frac{C}{B+C}$

V. Accuracy as for The Vigorous =  $\frac{A}{A+D}$

VI. Probability of Detection and Accuracy for The Depressed

$$= \text{II} \times \text{IV} = \frac{C}{C+D} \times \frac{C}{B+C}$$

Fig. 4. Definition of diagnostic probabilities in the prediction of neonatal conditions by a critical value of a single fetal variable at each stage of labor. (DRDN: Diagnostic Reliability for Depressed Neonates)

condition by this means are summed up in Fig. 5, as the number of cases in four categories, A, B, C and D which are shown in Fig. 4. A is the group of the correct diagnosis: "to be vigorous",

STAGE	CD=3-4-5 cm	CD=6-7 cm	CD=8-9 cm	early IInd	late IInd
A	=+ 78.854( )	=+ 808.32( )	=- 48.214( )	=- 62.068( )	=+ 3.8094(pH)
P	+ 0.050947( )	+ 0.21270( )	+ 7.0428( )	+ 2.2719( )	- 0.37797(pO <sub>2</sub> /10)
G	+ 4.6733( )	+ 36.132( )	+ 1.9029( )	+ 0.86729( )	+ 0.53180(pCO <sub>2</sub> /10)
A	- 1.0476( )	- 12.150( )	+ 2.0091( )	+ 0.93979( )	- 0.32041(BE)
R	- 2.6332( )	+ 1.1389( )	- 13.702( )	- 6.6476( )	+ 4.5486(BFHR/100)
	+ 0.18275( )	+ 2.2132( )	+ 2.7870( )	+ 0.31681( )	+ 0.27096(%UPI×10)
1	- 0.11629( )	- 0.94897( )	- 1.8864( )	- 0.96749( )	+ 0.17362(%CORD×10)
	-598.03	-6125.6	+360.30	+469.34	-41.074
A	=+ 23.344( )	=- 1610.0( )		=- 74.872( )	=+ 9.4500(pH)
P	- 0.94380( )	- 7.6174( )		+ 3.7620( )	- 1.4188(pO <sub>2</sub> /10)
G	+ 0.80012( )	- 90.465( )		- 0.27992( )	+ 1.1950(pCO <sub>2</sub> /10)
A	- 0.21998( )	+ 23.612( )		+ 0.63242( )	- 0.53702(BE)
R	- 5.00000( )	- 15.995( )		- 30.116( )	+ 3.9903(BFHR/100)
	+ 0.16011( )	- 0.87812( )		+ 0.59645( )	+ 0.042273(%UPI×10)
5	- 0.079156( )	+ 0.79376( )		- 1.1538( )	- 0.092590(%CORD×10)
	-166.24	+12298.0		+588.75	-82.039

Fig. 3. A pair of discriminant functions for the APGAR scores at 1 and 5 minutes after birth based upon a set of seven fetal parameters at a given stage of childbirth.

		CD=3-4-5cm	6-7 cm	8-9 cm	early II <sup>nd</sup>	late II <sup>nd</sup>
APGAR	Act.	11	3	0	1	10
	Pred.	8	3	3	5	11
1 min.						

		11	23
Act.		11	23
Pred.		8	5
1 min.			

		3	15
Act.		3	15
Pred.		3	1
1 min.			

		0	9
Act.		0	9
Pred.		3	0
1 min.			

		1	7
Act.		1	7
Pred.		5	1
1 min.			

		10	33
Act.		10	33
Pred.		11	4
1 min.			

		11	30
Act.		11	30
Pred.		3	3
5 min.			

		3	17
Act.		3	17
Pred.		2	0
5 min.			

Act.			
Pred.			
5 min.			

		1	11
Act.		1	11
Pred.		2	0
5 min.			

		5	48
Act.		5	48
Pred.		4	1
5 min.			

Fig. 5. The results of a number of cases as predicted by the set of discriminant functions as described in Fig. 3, utilizing seven fetal parameters, pH, PO<sub>2</sub>, PCO<sub>2</sub>, actual base excess, basal FHR, %UPI, and %CORD at a stage.

K: number of dependent variable  
 Y: dependent variable  
 X(i): independent variable  $i=1, \dots, K$   
 V(i,j): covariance matrix of X  
 COE(i): regression coefficient.

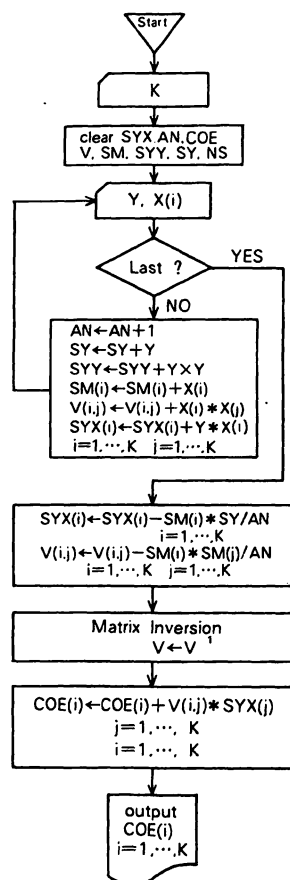


Fig. 6. Computing flow chart for partial regression analysis of APGAR scores at 1 and 5 minutes, respectively, using several fetal parameters.

and D is the newborns incorrectly diagnosed as "to be vigorous" although, actually, they were depressed.

## 2.2 Partial regression analysis

Fig. 6 is an example of a computing flow chart for obtaining a regression equation for APGAR scores at both one and five minutes after birth by partial regression analysis from the data at each stage of labor. The same sets of complete data of seven fetal parameters as in Tab. II were used for the regression in forty-seven cases at the early first stage. The equations, thus obtained, predicting the APGAR scores at both one and five minutes after birth are shown in Fig. 7. Some of the estimated APGAR scores are compared with actual ones in Tab. III.

The correlations between the actual APGAR score (ordinate) and the predicted one (abscissa) are plotted on graphs as shown in a pair of examples of Fig. 8. Then, the accuracy of the prediction was again tested by the four categories' representation as shown in the number of cases of Fig. 9.

## 2.3 Comparisons between the two computer techniques

According to the definition of probability of correct diagnosis (PCD) as  $(A + C)/(A + B +$

STAGE	CD=3-4.5 cm	CD=6-7 cm	CD=8-9 cm	early II nd	late II nd
<b>Apgar 1 =</b>					
	-57.190242(AI)	+84.920758(AII)	+26.101913(AIII)	-44.563751(AIV)	-40.465144(AV)
	+ 0.13237703(BI)	+ 0.11824057(BII)	+ 0.026806836(BIII)	+ 0.26802575(BIV)	+ 0.16502163(BV)
	- 0.34952788(CI)	+ 0.71923496(CII)	+ 0.13161483(CIII)	- 0.34332828(CIV)	- 0.21460333(CV)
	- 0.75636481(DI)	- 1.0496064(DII)	- 0.44784275(DIII)	+ 0.60535095(DIV)	+ 0.75342546(DV)
	+ 0.029290608(EI)	+ 0.014433495(EII)	- 0.012771252(EIII)	- 0.0035335822(EIV)	- 0.030099188(EV)
	- 0.59311087(FI)	- 2.8628641(FII)	- 3.5171248(FIII)	- 0.37129498(FIV)	- 2.0408312(FV)
	+ 0.90530631(GI)	+ 0.19212870(GII)	+ 2.2212651(GIII)	+ 3.5278252(GIV)	- 0.34620959(GV)
	+ 7.4680851	+ 7.6363636	+ 7.7500000	+ 6.6428571	+ 7.5000000
<b>Apgar 5 =</b>					
	-25.915916(AI)	+47.214978(AII)	+ 9.2879630(AIII)	+ 0.27197726(AIV)	-47.629169(AV)
	+ 0.034134722(BI)	+ 0.059117744(BII)	- 0.038835328(BIII)	+ 0.0043037027(BIV)	+ 0.053045685(BV)
	- 0.16556578(CI)	+ 0.35035301(CII)	+ 0.090798652(CIII)	- 0.069672665(CIV)	- 0.29220040(CV)
	+ 0.3456413(DI)	- 0.68235943(DII)	- 0.21765830(DIII)	+ 0.068284043(DIV)	+ 0.87152588(DV)
	+ 0.0054327945(EI)	+ 0.020527066(EII)	+ 0.0075318463(EIII)	+ 0.0056399830(EIV)	- 0.023061280(EV)
	- 0.94582987(FI)	- 1.6034803(FII)	- 2.0152945(FIII)	- 0.38603877(FIV)	- 0.059476791(FV)
	+ 0.66506851(GI)	+ 0.42703847(GII)	+ 0.40882846(GIII)	+ 1.2457307(GIV)	+ 0.19240421(GV)
	+ 8.5744680	+ 8.6818181	+ 8.9166666	+ 8.6428571	+ 8.6379310
	AI=pH-7.2902123	AII=pH-7.2813630	AIII=pH-7.2683328	AIV=pH-7.2814281	AV=pH-7.2646548
	BI=pO <sub>2</sub> -18.925529	BII=pO <sub>2</sub> -17.677270	BIII=pO <sub>2</sub> -16.958332	BIV=pO <sub>2</sub> -16.621427	BV=pO <sub>2</sub> -15.086203
	CI=pCO <sub>2</sub> -47.972336	CII=pCO <sub>2</sub> -46.738358	CIII=pCO <sub>2</sub> -47.316663	CIV=pCO <sub>2</sub> -47.114280	CV=pCO <sub>2</sub> -50.112064
	DI=BE+4.1234036	DII=BE+5.1227269	DIII=BE+6.1083330	DIV=BE+5.4642854	DV=BE+5.1931031
	EI=BFHR-145.63829	EII=BFHR-152.500000	EIII=BFHR-154.16666	EIV=BFHR-147.14285	EV=BFHR-143.62068
	FI=%UPI-0.21163995	FII=%UPI-0.34818840	FIII=%UPI-0.31627886	FIV=%UPI-0.41656874	FV=%UPI-0.46783039
	GI=%CORD-0.38668012	GII=%CORD-0.43944969	GIII=%CORD-0.52043650	GIV=%CORD-0.47154710	GV=%CORD-0.67467240

Fig. 7. A pair of partial regression functions for 1 and 5 minutes' APGAR scores based upon a set of seven fetal parameters at a given stage of childbirth.

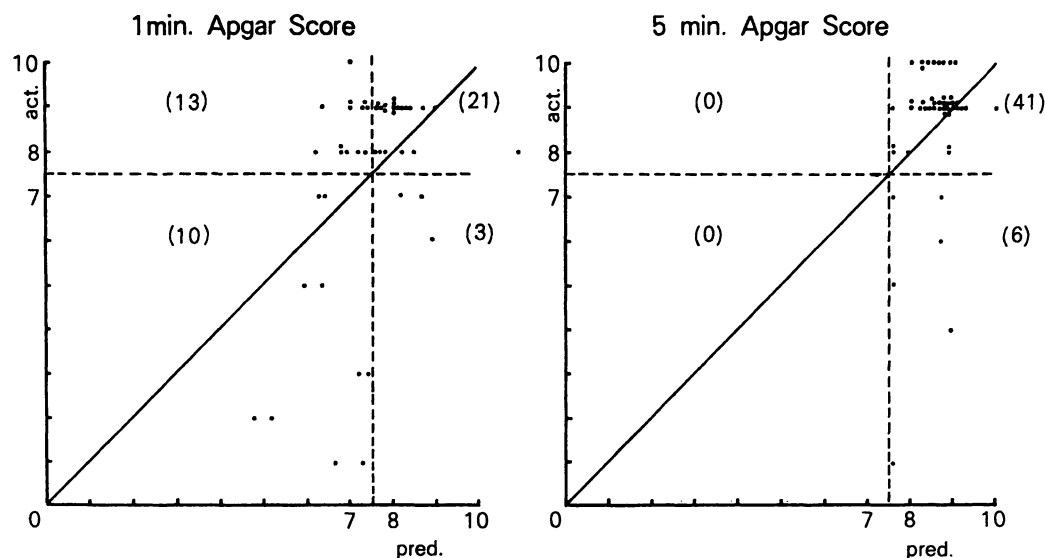


Fig. 8. A pair of examples of a correlogram between the estimation by the two partial regression functions at the early first stage of labor and the actual APGAR cores at 1 and 5 minutes after birth.

	CD=3-4.5cm	6-7cm	8-9cm	early II nd	late II nd
A P G A R	Act.	Act.	Act.	Act.	Act.
	13	5	1	3	13
1 min.	Pred.	Pred.	Pred.	Pred.	Pred.
	21	13	8	5	30
A P G A R	Act.	Act.	Act.	Act.	Act.
	10	4	3	6	11
5 min.	Pred.	Pred.	Pred.	Pred.	Pred.
	3	0	0	0	4
A P G A R	Act.	Act.	Act.	Act.	Act.
	0	1	0	0	3
5 min.	Pred.	Pred.	Pred.	Pred.	Pred.
	41	19	12	12	50
A P G A R	Act.	Act.	Act.	Act.	Act.
	0	1	0	0	2
5 min.	Pred.	Pred.	Pred.	Pred.	Pred.
	6	1	0	2	3

Fig. 9. Prediction of APGAR scores with partial regression functions as described in Fig. 7, using the seven fetal parameters pH, PO<sub>2</sub>, PCO<sub>2</sub>, actual BE, mean BFHR, %UPI, and %CORD at a stage.

Tab. III. Examples of predicted APGAR scores by partial regression analysis.

Case No.	Actual APGAR s.	Estimated APGAR s.
60	5.00000	6.38989
61	7.00000	8.65149
62	9.00000	7.38468
64	5.00000	5.97247
71	9.00000	7.40897
72	8.00000	6.92426
73	8.00000	8.21070
74	9.00000	7.02310
75	6.00000	8.91809
78	8.00000	7.35747
80	3.00000	7.45556
81	2.00000	4.84235
82	9.00000	8.13095
83	9.00000	8.78852
84	8.00000	6.81620
88	3.00000	7.24268
92	9.00000	8.27731
93	8.00000	8.49912
96	9.00000	8.01982
97	9.00000	7.83361
98	9.00000	8.06337
99	7.00000	8.22959
100	9.00000	8.03133

C + D) in the four groups shown in Fig. 4, the diagnostic probabilities of these two techniques at each stage are compared with APGAR scores at one and five minutes. The diagnostic reliability for the depressed neonates (DRDN), which is defined as  $C/(C + D) \times C/(B + D)$  is the prob-

ability to satisfy necessary as well as sufficient conditions for the diagnosis of the depressed infants, is also compared. Both the above, PCD and DRDN, are calculated from the numbers in Fig. 5 and Fig. 9 and are summarized in Tab. 4.

### 3. Comment

The computer, often called an artificial brain, has opened a new era in techniques, replacing the human brain by electronic logic and memory devices. Astronomical, weather and economical trend predictions are based upon certain logical processing of up-to-date information which may be vastly complicated and yet analyzable by statistical means. Prediction of fetal outcome at birth is quite a similar situation although it is also very complicated and it is difficult to believe in the continuity of the fetus in utero to the extrauterine neonate. Too many unpredictable interferences can occur during such a short period of time. It is hardly possible to imagine that such a dynamic unpredictability of childbirth can be overcome by logical or rational judgments using various kinds of measurements or information. How much is contributed to this problem by computer techniques. The preliminary results in this report are seemingly not so satisfying. If the probability of correct diagnosis (PCD) by computer techniques is compared with that by the simple pH criterion [2, 6, 7] in the same group of subjects, the percentages are not so

Tab. IV. Comparison between diagnostic probability—I (PCD) and—VI (DRDN) for correct discrimination of neonatal conditions in four categories' diagnoses by a pair of computer predictions (DFA and PRA) based upon 7 fetal parameters (Fig. 5 und Fig. 9) and that of a single fetal criterion that if the pH is less than 7.20 the infant will be depressed (Fig. 10).

Stage		CD=3—4—5 cm		CD = 6—7 cm		CD = 8—9 cm		Early IInd		Late IInd	
APGAR		1 min	5 min	1 min	5 min	1 min	5 min	1 min	5 min	1 min	5 min
(D) (F) (A)	PCD%	66.0	70.2	81.8	86.4	100.0	—	85.7	86.7	75.9	89.7
	DRDN%	25.9	10.7	37.5	40.0	100.0	—	69.5	66.7	28.4	35.6
(P) (R) (A)	PCD%	66.0	87.3	77.3	90.9	91.7	100.0	78.6	85.7	70.7	89.7
	DRDN%	33.5	0.0	44.4	25.0	75.0	0.0	66.7	0.0	33.6	16.0
(pH)	PCD%	70.2	85.1	86.4	95.5	75.0	66.7	64.3	78.6	75.8	82.8
	DRDN%	0.0	0.0	25.0	50.0	33.3	0.0	16.7	0.0	18.5	8.9

DFA: discrimination function analysis, PRA: partial regression analysis

PCD: probability for correct diagnosis  $(A + C)/(A + B + C + D)$

DRDN: diagnostic reliability for depressed neonates  $C/(C + D) \times C/(B + C)$

		CD : 3-4-5cm		6-7cm		8-9cm		early II nd		late II nd	
A P G A R	Act.	1	33	0	18	2	7	0	8	4	39
	1 min.	0	13	1	3	2	1	1	5	5	10
		Pred.									
A P G A R	Act.	1	40	0	20	0	8	1	11	7	46
	5 min.	0	6	1	1	0	4	0	2	2	3
		Pred.									

Fig. 10. Discrimination of neonatal conditions by a single fetal pH criterion, i. e. if the fetal capillary blood's pH is lower than 7.20, then the infant is expected to be depressed. The subjects are the same as computed for Fig. 5 and Fig. 9, as described in Tab. 1.

much higher in the former than the latter even after such complicated processing of the many parameters (Fig. 10 and Tab. IV). The possible reasons for this are as follows:

1. The probability for correct diagnosis (PCD) is quite dependent upon the percentage incidence of vigorous babies among the total number of cases, and it is fairly difficult to achieve a marked increase in it by any means.
2. The number of the cases fed into the computer is too small and the distribution of the data deviates too much from the GAUSSIAN distribution to have a good resolution.
3. The processed data sampled during twenty minutes at a given stage are not sufficient enough to clarify the pathogenesis of fetal and neonatal depression. For example, during the sampling period of twenty minutes acute depression of the fetus due to interrupted umbilical circulation [5] just several minutes before birth cannot always be reflected in the averaged value of fetal parameters. In fact about one third of the cases of neonatal depression in our series can be ascribed to such acute depression.

Sampled data have definite shortcomings. Sampling of fetal blood can never be done continuously in present circumstances, so that the continuous monitoring of FHR is quite informative in that sense. However, even with such sampled information in the blood the fetal

**pH has great diagnostic significance better than any other parameters** as seen in the high achievement in PCD (DP-I) by the single pH criterion of Tab. IV. The fact that the largest coefficient was found in the pH term in all the equations of Fig. 3 and Fig. 7 is also quite convincing. Therefore, sampled data of fetal blood may never be considered as perfect information but are still quite important for clinical judgment. **The continuous data provided by FHR monitoring is not a decisive factor but rather a screening aid** to complement the decision when, and in what case, to take a fetal blood specimen from the presenting part. In that sense the sampled data system described here was represented by an averaged value of each parameter during a sampling period of twenty minutes. This timing was rather random because a stage of labor for data sampling was defined according to the cervical dilatation. In other aspects, however, it is a clinically meaningful situation because of the fact that it is not always possible to monitor all childbirths from the beginning to the end. In practice obstetricians frequently face the situation where they have to make a decision to deliver a fetus, through the vagina or by cesarean section, just on a limited amount of information.

Sometimes it is rather dangerous to sit at a monotonous monitoring desk, because we fre-



quently become accustomed to seeing slight abnormal findings and hence we lose a sense of good timing that would allow us a decisive action. Analog data are easily overlooked even with the slightest but definite abnormalities, so that quantification of the analog signals should be an obligatory process for human pattern recognition. Also the digitized data loses some information, if it is done by certain limited number of samplings. Both the analog and the digital data processings should be useful. In fact as described previously, the continuous monitoring of FHR is the way to detect fetal distress at the earliest possible time. This produces, however, many false positive results in otherwise vigorous infants. On the contrary, the pH of fetal blood is quite a decisive parameter in predicting a definitely depressed infant if it is truly lower than normal, although it is rather late to have a real decrease of pH value in the fetal blood.

If we turn our eyes from the task of screening normal babies out of the depressed who should be detected as early and as accurately as possible, it is astonishing that the diagnostic reliability for the depressed neonate (DRDN) in Tab. IV is very much improved by both computer techniques, surpassing the results of single parameter prediction by means of fetal pH. The various merits of seven fetal parameters all seem to be contributing to these results, which are encouraging enough for us to revise the mathematical models upon which we have processed the data. More extensive data should be accu-

mulated from a larger population. The type of distribution of data and the correlation between them should be rechecked. The time course of each fetal parameter during the five stages in utero can be informative in the prediction, even though it is unrealistic to expect to get a full set of information at every stage of labor, in all individuals. Hence, if we can add some defective data by statistically reasonable interpolation by a computer we shall be able to produce a marked increase in the number of processable cases with better prediction in the future.

#### 4. Conclusions

1. Two computer prediction systems of fetal outcome in terms of 1 and 5 min. APGAR scores based upon complete sets of seven fetal parameters i. e., pH,  $PO_2$ ,  $PCO_2$ , actual BE, mean BFHR, %UPI and %CORD which were sampled at each stage of labor (early, middle, late 1st stage, early and late 2nd stage) have been developed, using the data of biochemically and electronically monitored childbirths at the Yale New Haven Medical Center.

2. Neither computer techniques, discriminant analysis or partial regression analysis, could achieve a marked increase in the PCD (probability for correct diagnosis) over the results of a single pH criterion, but they did greatly improve the DRDN (diagnostic reliability for the depressed neonate), which encourages us in the further automatic monitoring of childbirths.

#### Summary

In order to avoid subjectively biased judgment, two computer prediction systems for 1 and 5 minutes' APGAR scores have been developed on the basis of multi-variable analysis of fetal data. One uses discriminant analysis to differentiate the depressed infant of APGAR 7 or less from the normal of 8 or more, the other is partial regression analysis to estimate the APGAR score itself by a regression function derived from a set of fetal variables at each stage of labor. More than one hundred cases were monitored continuously by electronic as well as biochemical means at Yale New Haven Medical Center in 1966-67, but only cases with complete sets of information (such as fetal capillary blood pH,  $PO_2$ ,  $PCO_2$ , actual base excess (BE), the level of basal fetal heart rate (BFHR) and the percentage incidences of late and variable FHR decelerations over a number of uterine contractions (%UPI and %CORD), during a sampling period of twenty minutes at each stage of

labor, were provided for computation as shown in Tab. I. A course of labor was divided into five stages, namely the early first stage of 3-4-5 cm cervical dilation, the middle first stage of 6-7 cm, the late first stage of 8-9 cm, the early second stage, and the late second stage which was the last twenty minutes just before delivery. Several times' measurements of fetal blood pH were made with an ASTRUP pH-meter. The  $PO_2$  and  $PCO_2$  were measured by a micro IL meter during the sampling period of twenty minutes and were averaged to give a set of data for one case at each stage. Base excess was calculated from the SIGGAARD-ANDERSEN alignment nomogram. The level of the BFHR was also an averaged value of the basal FHR during all uterine contractions during the twenty minutes of continuous recording. Such sampled data from the fetus is not complete monitoring but is, nevertheless, in practice, quite important, because the obstetrician often has

to make a decision as to deliver a fetus vaginally or abdominally on a relatively limited amount of information. Some examples of such data for computation are shown in Tab. II. The respective discriminant functions for one and five minutes' Apgar scores, at any one of the five stages, are summarized in Fig. 3 so as to give a positive value if the fetus is to be born vigorous and a negative one if depressed.

According to correlograms between the computer prediction and the actual outcome shown in Fig. 4, the number of cases of correct or false diagnosis are summed up in Fig. 5. The probability for correct diagnosis [ $PCD = (A + C)/(A + B + C + D)$  in Fig. 4] is about 85 per cent as shown in Tab. IV.

Another computation was carried out to get partial regression functions (Fig. 7) in order to predict one and five minutes' Apgar scores themselves, as shown by the examples in Tab. 3. The probability for correct diagnosis by these equations is once again evaluated by the number of cases of correct and false discrimination and shows 70–90% accuracy and is therefore similar to the results obtained by discriminant analysis. However, as Tab. IV clearly shows, these probabilities for correct diagnosis are little better than the achievement of the simplest, single pH discrimination, which states that if the fetal capillary blood pH is greater than or equal to 7.20, the baby should not be depressed. This poor result, even by computation, can be ascribed to several factors:

1. The percentage of vigorous infants among all the cases studied was too high to achieve a marked increase in the probability for correct diagnosis (PCD).
2. The number of cases was too small especially in the depressed and deviates too much from the Gaussian distribution.
3. The sampled and averaged data in a twenty minute period of a given stage did not provide precise enough information.

However, these methods demonstrate the great significance of the fetal blood pH. This is shown by the pH term, which has the largest coefficient value in any of the equations not only of the discriminant functions but also of the partial regression functions. On the other hand when we consider the diagnostic reliability for the depressed neonate [ $DRDN = C/(C + D) \times C/(B + C)$  in Fig. 4], it is surprising that the two computer techniques have made such improved probabilities as high as 30–70% in contrast to the single pH judgments which have, on average, a DRDN of about 25% (including many zeros at various stages, Tab. IV). These differences should mean great improvement in the clinical management of very sick fetuses, because detection of the depressed infants as early as possible and as accurately as possible is the major task of fetal diagnosis during labor. The various merits of seven fetal parameters all seem to be contributing to these improved results. Hence we are sufficiently encouraged to accumulate more data and to revise our mathematical models upon which we have processed the data.

**Keywords:** Computer prediction, Apgar score, fetal heart rate, fetal capillary blood, discriminant analysis, partial regression analysis, diagnostic probabilities.

## Zusammenfassung

### Computervorhersage des Neugeborenenzustandes

Um subjektive Einflüsse auszuschließen, wurden auf der Grundlage einer Vielfaktorenanalyse der fetalen Meßdaten zwei Computer-Vorhersagesysteme für den 1- und 5-Minuten Apgar-Wert entwickelt. Das eine benützt die Diskriminanzanalyse zur Unterscheidung des deprimierten Kindes mit einem Apgar von 7 oder weniger von dem normalen mit einem Apgar von 8 oder mehr. Das andere benützt die partielle Regressionsanalyse, wobei der Apgar-Wert selbst geschätzt wird, als Regression einer Funktion, die sich von einem ganzen Satz fetaler Variabler in den einzelnen Stadien der Geburt ableitet. Über 100 Fälle wurden am Yale New Haven Medical Center in der Zeit von 1966–1967 kontinuierlich sowohl elektronisch als auch biochemisch überwacht. Aber nur Fälle mit einem vollständigen Satz an Informationen (nämlich pH,  $PO_2$ ,  $PCO_2$  und aktueller Basen Exzess im fetalen Kapillar-Blut, die basale Herzfrequenz des Feten, und prozentuale Angaben über Spättiefs und variable Tiefs der fetalen Herztöne bezogen auf die Zahl der Wehen) und zwar über eine 20-Minuten-Phase für jedes Stadium der Geburt wurden für die Computer-Auswertung verwendet wie dies in Tab. I gezeigt wird. Der Geburtsverlauf wurde in 5 Phasen eingeteilt, nämlich die frühe Eröffnungsperiode bei einer Zervixdilatation von 3–4–5 cm, die mittlere Eröffnungsperiode bei einer Zervixdilatation von 6–7 cm, die späte

Eröffnungsperiode bei 8–9 cm, die frühe und die späte Austreibungsperiode; letztere umfaßte die letzten 20 Minuten vor der Entbindung. Mehrfach wurden Messungen des fetalen Blut-pH mit einem ASTRUP-Gerät durchgeführt.  $PO_2$  und  $PCO_2$  wurden mit einem Micro-IL-Meßgerät während der 20-Minuten-Phasen gemessen, und ein Mittelwert errechnet, um für jeden Fall zu jedem Stadium einen Satz an Meßdaten zu erhalten. Der Basen-Exzess wurde nach dem SIGGAARD-ANDERSEN-Nomogramm ermittelt. Die Höhe der Basalherzfrequenz des Feten war ebenfalls ein gemittelter Wert der basalen fetalen Herzfrequenz während der Wehen innerhalb der 20-Minuten-Phasen, die kontinuierlich registriert wurden. Die so stichprobenartig gewonnenen Meßdaten des Feten bedeuten keine vollständige kontinuierliche Überwachung, sind aber dennoch in der Praxis von großer Bedeutung, da der Geburtshelfer oft mit einer relativ begrenzten Anzahl von Informationen entscheiden muß, ob die Entbindung auf vaginalem oder abdominalem Wege zu Ende zu führen ist. Einige Beispiele solcher Meßdaten für die Computerauswertung sind in Tab. II dargestellt. Die entsprechenden Diskriminanz-Funktionen für den 1-Minuten und 5-Minuten Apgar-Wert während jeder der 5 Geburtsphasen sind in Fig. 3 zusammengestellt, wobei sich ein positiver Wert ergibt, wenn der Fetus in einem lebensfrischen und ein negativer, wenn er in einem deprimierten Zustand geboren werden

wird. Entsprechend der Gegenüberstellung der Computervorhersage und dem tatsächlichen Ergebnis wie es in Fig. 4 dargestellt ist, sind die Fallzahlen der korrekten und der falschen Diagnosen in Fig. 5 zusammengestellt. Die Wahrscheinlichkeit der richtigen Diagnose liegt bei etwa 85% wie es in der Tab. IV gezeigt wird. Ein anderes Computerauswertungssystem wurde zur Ermittlung partieller Regressionsfunktionen verwendet um die 1-Minuten und 5-Minuten-APGAR-Werte selber vorherzusagen, wie anhand von Beispielen in der Tab. III dargestellt ist. Die Wahrscheinlichkeit der richtigen Diagnose bei diesem Rechengang wurde ebenfalls durch die Gegenüberstellung der Zahl der Fälle richtiger und falscher Einschätzung ermittelt; sie zeigt eine Genauigkeit von 70–90% und ist damit vergleichbar mit den Ergebnissen die durch die Diskriminanzanalyse erzielt wurden. Die Wahrscheinlichkeit der richtigen Diagnose aber ist, wie Tab. IV klar darlegt, größer als wenn man sie von einer einfachen einzelnen pH-Bestimmung ableiten würde, was die Feststellung zuläßt, daß bei einem fetalen kapillaren Blut-pH von 7,20 und höher das Neugeborene nicht deprimiert sein wird.

Dieses spärliche Ergebnis der Computerauswertung kann verschiedenen Faktoren zugeschrieben werden:

1. Der Prozentsatz der lebensfrischen Kinder unter den untersuchten Fällen war zu hoch, um einen deutlichen Anstieg über den Normalprozentsatz zu erzielen.
2. Die Fallzahl war zu klein und streut zu sehr über die GAUSSsche Verteilung hinaus.

**Schlüsselwörter:** Computervorhersage, APGAR-Wert, fetale Herzfrequenz, fetales Kapillarblut, Diskriminanzanalyse, partielle Regressionsanalyse, diagnostische Wahrscheinlichkeiten.

## Résumé

### Prédiction par ordinateur de l'état du nouveau-né grâce à l'analyse d'informations sur le rythme cardiaque foetal et le sang capillaire au cours du travail

Afin d'éviter les jugements subjectifs et partiels, deux systèmes de prédiction, par ordinateur, ont été établis pour les index d'APGAR à 1 et 5 minutes par des méthodes d'analyse de variables multiples à partir d'informations objectives sur le fœtus.

L'une est une analyse discriminante permettant de distinguer les nouveau-nés déprimés avec un APGAR égal ou inférieur à 7 de ceux qui sont normaux avec un APGAR égal ou supérieur à 8. L'autre est une analyse de régression partielle permettant d'évaluer l'index d'APGAR lui-même par une fonction de régression calculée à partir d'un ensemble de variables fœtales à chaque stade du travail.

Parmi les cent cas avec surveillance continue par moyens électroniques et biochimiques au Centre Hospitalier de l'Université de Yale en 1966 et 1967, seuls les cas avec un ensemble d'informations complet (pH,  $PO_2$ ,  $PCO_2$ , excès de base (B. E.), niveau de rythme cardiaque foetal de base (B. F. H. R.), et pourcentage de décélérations tradives et variables par rapport au nombre de contractions utérines (%UPI et %CORD) pendant une période de vingt minutes au cours du travail, ont été utilisés pour les calculs comme

3. Die innerhalb von 20-Minuten-Phasen für jede Geburtsphase gewonnenen, und gemittelten Meßdaten lieferten nicht genügend Informationen. Dennoch zeigen diese Methoden die große Bedeutung des fetalen Blut-pH. Dies ist dargestellt am pH-Wert, der den größten Koeffizienten in allen Rechengängen hat, nicht nur bei der Diskriminanz-Funktion sondern auch bei der partiellen Regressionsfunktion.

Wenn wir andererseits die diagnostische Zuverlässigkeit bezüglich des deprimierten Neugeborenen betrachten, so ist es verwunderlich, daß die beiden Computertechniken die Wahrscheinlichkeit auf 30–70% erhöhten. Demgegenüber beträgt die Wahrscheinlichkeit einer korrekten Vorhersage bei alleiniger Beurteilung durch eine Einzel-pH-Messung durchschnittlich etwa 25% (Tab. IV). Diese Unterschiede können eine bedeutende Verbesserung bedeuten für die klinische Betreuung sehr kranker Feten, denn die möglichst frühe und möglichst genaue Erkennung des deprimierten Kindes ist die vorrangige Aufgabe der fetalen Diagnostik unter der Geburt. Die sieben verschiedensten fetalen Parameter scheinen alle zu diesen verbesserten Ergebnissen beizutragen. Von daher sind wir ausreichend ermutigt, weiteres Material zusammenzutragen und unsere mathematischen Modelle zu überarbeiten, mit denen wir die Meßdaten ausgewertet haben.

le montre le Tab. I. Le travail a été divisé en cinq stades: dilatation cervicale de 3–4–5 cm, 6–7 cm, 8–9 cm, stade de début d'expulsion et les 20 minutes précédant l'accouchement. Si plusieurs mesures du pH sanguin foetal par l'appareil d'ASTRUP et de la  $PO_2$  et la  $PCO_2$  par une méthode de micro-mesure étaient pratiquées dans chaque période de 20 minutes, on faisait la moyenne des résultats obtenus pour avoir un ensemble de données à chaque stade.

L'excès de base a été calculé sur le graphique de SIGGAARD-ANDERSEN. Le rythme cardiaque foetal de base a été obtenu en faisant la moyenne des rythmes de base (B. F. H. R.) pendant 20 minutes. Les informations obtenues sur le fœtus ne suffisent pas pour réaliser une surveillance complète. Elles sont cependant tout-à-fait importantes pour aider à porter un jugement en clinique parce qu'en pratique l'obstétricien doit souvent prendre la décision d'extraire le fœtus par voie vaginale ou abdominale sur un nombre limité d'informations.

Le Tab. II montre des exemples de données utilisables pour les calculs. Dans la première étude, les fonctions discriminantes ont été conçues de telle sorte qu'on obtienne un chiffre positif si le nouveau-né est vigoureux, et négatif s'il est déprimé comme on l'a résumé Fig. 3 pour chacun des cinq stades et pour les index d'APGAR une minute et cinq minutes respectivement.

Les quatre catégories différentes du tableau de corrélation entre les valeurs prédites et réelles sont montrées Fig. 4. Le nombre de cas avec diagnostic exact ou faux est présenté Fig. 5. La probabilité d'un diagnostic exact ( $PCD = \frac{A + C}{A + B + C + D}$ ), c'est-à-dire le nombre de cas dans lesquels l'état du fœtus (déprimé ou non) est prédit correctement est de l'ordre de 85% en moyenne avec les limites de 66% et 100% (Tab. IV).

On a utilisé le même ensemble de données pour un autre calcul, celui des fonctions de régression partielle (Fig. 7) afin de prédire les index d'APGAR 1 et 5 minutes à chaque stade du travail sous forme digitale, tels qu'on le montre Tab. III. Les probabilités de donner un diagnostic exact (P. C. D.) à partir de ces équations sont calculées pour les patientes ayant eu par l'étude discriminante des prédictions exactes ou erronées. Le calcul donne des résultats identiques, pour 70 à 90% à ceux de l'étude discriminante (Tab. 4).

Cependant, comme il est démontré clairement tab. 4, ces probabilités de diagnostic exact ne sont guère meilleures que celles obtenues par l'emploi du pH seulement, ce qui signifie que si le pH du sang capillaire fœtal est supérieur ou égal à 7,20, le nouveau-né ne devrait pas être déprimé et vice-versa. Ce résultat peut être imputé à plusieurs facteurs:

- 1°) Le pourcentage d'enfants vigoureux est trop élevé pour qu'on puisse atteindre un niveau supérieur,
- 2°) Le nombre de patientes est trop petit et trop dévié par rapport à la distribution de GAUSS, et,

- 3°) Les données que l'on a mesurées et dont on a fait la moyenne en 20 minutes pour un certain stade ne peuvent pas apporter suffisamment d'informations.

Mais par n'importe quelle voie on peut démontrer la haute valeur diagnostique du pH sanguin fœtal même par le coefficient élevé obtenu pour le pH dans chacune des équations, non seulement dans les fonctions discriminantes mais aussi dans les fonctions de régression partielle (Fig. 3 et 7).

Par ailleurs, si nous examinons la fiabilité du diagnostic ( $DRDN = \frac{C}{C + D} \times \frac{C}{B + C}$ ) dans le cas de nouveau-nés déprimés, en ce qui concerne les conditions nécessaires mais aussi suffisantes pour détecter l'état de dépression aussitôt que possible et aussi exactement que possible, il est surprenant d'avoir une élévation des probabilités de l'ordre de 30 à 70% environ par les deux méthodes de calcul sur le simple pH qui représente environ 25% de probabilités mais avec nombreux zéros à différents stades (Tab. 4). Ces différences pourraient se traduire par une importante amélioration dans l'attitude pratique, en Clinique, pour les fœtus en très mauvais état, puisque c'est chez ces derniers que se posent les problèmes diagnostiques majeurs. Les différentes valeurs des sept paramètres fœtaux semblent toutes jouer un rôle dans les résultats, qui sont assez encourageants pour que d'autres données soient réunies et que l'on tente de vérifier les méthodes de calcul utilisées pour le traitement de nos informations.

**Mots-clés:** Prédiction par ordinateur, index d'APGAR, rythme cardiaque fœtal, sang capillaire fœtal, analyse discriminante, analyse par régression partielle, probabilités diagnostiques.

### Acknowledgements

The clinical parts of this work were carried out when one of the authors was at Yale University School of Medicine as a FULBRIGHT research fellow from the U. S. Department of State and as a JAMES-HUDSON-BROWN Memorial postdoctoral fellow from Yale during 1966—67. The authors express sincerest thanks to Prof. K. KURACHI of Osaka University Medical School and to Prof. E. J. QUILLIGAN and Dr. E. H. HON of University Southern California for their consistent encouragement of this work.

### Bibliography

- [1] APGAR, V.: The newborn (APGAR) scoring system. *Pediatric Clinic of North America* 13 (1966) 645
- [2] BEARD, R. W., E. D. MORRIS: Fetal and maternal acid-base balance during normal labour. *J. Obstet. Gynaec. Brit. Cwlth.* 72 (1965) 496
- [3] DRÄGE, T. S., C. KENNEDY, H. BERENDES, B. K. SCHWARZ, W. WEISS: The APGAR score as an index of infant morbidity. *Develop. Med. Child. Neurol.* 8 (1966) 141
- [4] HON, E. H.: Detection of fetal distress. Fifth World Congress of Gynaecology and Obstetrics. In: WOOD, C.: Fifth World Congress of Gynaec. & Obstet. Butterworths, Sydney 1967
- [5] HON, E. H., A. F. KHAZIN: Biochemical studies of the fetus. *Obst. & Gynec.* 33 (1969) 219
- [6] KUBLI, F.: *Fetale Gefahrenzustände und ihre Diagnose.* Thieme, Stuttgart 1966
- [7] SALING, E.: *Das Kind im Bereich der Geburtshilfe.* Thieme, Stuttgart 1966
- [8] SIGGAARD-ANDERSEN, O.: *The Acid-base Status of the Blood.* Williams & Wilkins, Baltimore 1966
- [9] TAKEMURA, H.: Optimization in discriminating prediction of APGAR scores by a fixed criterion of fetal parameters, pH,  $PO_2$ ,  $PCO_2$ , BE, actual BE, basal FHR, %UPI, and %CORD (submitted for publication)

Hikaru Takemura, M. D.  
Department of Obstetrics & Gynecology  
Osaka University Medical School  
1—2, Dojimahamadori 3-chome  
Fukushima-ku, Osaka/Japan